

5th International Conference on Corpus Linguistics (CILC2013)

Possible Effects of Free on Line Data Driven Lexicographic Instruments on Foreign Language Learning: The Case of *Linguee* and the *Interactive Language Toolbox*

Kris Buyse*, Serge Verlinde

Leuven Language Institute (KU Leuven, University of Leuven), Dekenstraat 6, 3000 Leuven, Belgium

Abstract

In this article we will examine whether certain data driven dictionaries can improve user-friendliness and usability –in a broader sense- of lexicographic instruments used in language learning, as well as enhance positive features such as empowerment, discovery, learner autonomy and the correct identification of lexico-grammatical patterns. Therefore we present the qualitative and quantitative data of a small experiment with the on line multilingual (English-Spanish) tool *Linguee*, comparing its efficiency and satisfaction levels at a Spanish writing test and the ones of the same tests in which, in one control group, only traditional dictionaries were used, and in another those dictionaries in combination with *Linguee*. Based on our findings, we will propose some further pedagogical and technological interventions in order to obtain even higher satisfaction and efficiency levels. An example of a result of such interventions can be found partially in the Interactive Language Toolbox.

© 2013 The Authors. Published by Elsevier Ltd.
Selection and peer-review under responsibility of CILC2013.

Keywords: lexicography, dictionaries, corpora, data driven, foreign language teaching

1. Introduction

Dictionaries are on the bookshelf of any student, as they seem to be the perfect tool for enhancing learner autonomy, one of the purposes of modern language teaching. Nevertheless, they are not used sufficiently nor adequately due to factors such as time, insecurity, lack of knowledge of the possibilities offered by the dictionary.

* Corresponding author. Tel.: +32 16 32 47 68; fax: +32 16 32 47 67.
E-mail address: kris.buyse@arts.kuleuven.be

Moreover, the information contained in the dictionaries can be limited, inadequate or lacking context (Nomdedeu, 2009).

Today, given the current state of technology, it is useful to investigate which technical and content-related innovations we can reasonably expect to find in a modern electronic data driven dictionary (or “dictionary-featured corpora”?), with a particular focus on the access to lexicographic description and the interaction between user and dictionary.

In this study our research question is whether such data driven dictionaries can improve user-friendliness and usability, as well as enhance further more positive features such as empowerment, discovery, learner autonomy and the correct identification of lexico-grammatical patterns. Therefore, after outlining the current context of teaching (Spanish) writing (2), we will present the qualitative and quantitative data of a small experiment with such a dictionary, i.e. *Linguee* (www.linguee.com), in comparison with more “traditional” tools (3). We will conclude by proposing, based on our findings, additional pedagogical and technological interventions in order to obtain even higher satisfaction and efficiency levels (4).

2. Teaching (Spanish) writing to the i-Gen: new generations, new tools

Language learners of Spanish use to burst out laughing when they see “bachelor in Law” translated by a machine as “soltero en derecho”, but their laughter turns into incomprehension when they find that the same kind of mistakes are made by students themselves, as –in the same way as a computer algorithm– they tend to copy the first dictionary translation and paste it without taking into account the paradigmatic and syntagmatic conditions imposed by the context.

2.1. Teaching (Spanish) writing to the i-Gen

As Buyse (2011) already pointed out, the internet-Generation (“i-Gen”), educated with less authoritarian methods than his predecessors, does not—at least not to the extent that the teacher and educational authorities wish he would— take responsibility and initiatives. Neither does the student make use of the instruments presented by the teacher. As a result, the student's progress is minimal, unless triggers and rewards are provided (Gyselaers, 2007).

However, psycholinguistic research on learning effects reports that 80% of the structures learnt are lost in less than 24 hours (Cervero & Pichardo Castro, 2000: 130), and that “deep level processing”, i.e. recurrent exposure to the material in different contexts, is the only remedy (Schmitt, 2000: 129).

When assessing students’ competences, teachers often complain that learners’ factual competence contrasts sharply with their linguistic competence (Nielsen & Mourier, 2007: 123). Our own analysis of the results of the level tests at the beginning of our Spanish courses (Buyse, 2008) indicates that students’ attention to lexical, grammatical and orthographic accuracy is limited. The results of the tests clearly show that this is not a very effective way of acquiring vocabulary, as word-by-word memorization does not provide learners with the necessary contextualisation, terminological variation and useful word combinations. On the other hand, students tend to translate literally from their mother tongue (Bowker & Pearson, 2002: 18).

2.2. New generations, new tools: first results

As a result, it is essential that students make (correct and frequent) use of certain instruments enabling them to avoid a range of basic and frequent problems, since users continue to encounter unforeseen problems, even in simple searches (Pastor & Alcina, 2010). In other words, requesting assistance of those whom we call metaphorically “the 6 experts”, namely spelling checkers, dictionaries, corpora, grammars, native speakers and teachers (Buyse, Delbecque & Speelman, 2009).

Data of Buyse (2011, 391) suggest a strong impact of the use of writing portfolios in which students have to document the use of instruments during the writing process on the overall scores, especially the influence of use of spelling checker on orthography scores and of digital dictionaries on vocabulary.

and pragmatics. A qualitative analysis was performed on the answers of the students on a brief questionnaire about their overall satisfaction on the quality of the instruments as a whole and of Linguee in particular, as well as on the time available for the experiment, among other items.

Our research questions were the following: can a data-driven tool such as Linguee (1) improve user-friendliness and usability and (2) enhance better scores, especially for lexico-grammatical patterns. Our research hypothesis was that CG 2 would score best (both on the quantitative and on the qualitative tests) because of the richness of the tools and its impact on the linguistic accuracy and the satisfaction on the quality of the tools.

Surprisingly, comparing to the medians of tests with earlier student populations of the same course of study (2009-2011), EG (with only Linguee at their disposal) scores best, except for orthography (see Figure 2), and the difference with the other groups is particularly high (2 points and more) on the vocabulary scores.

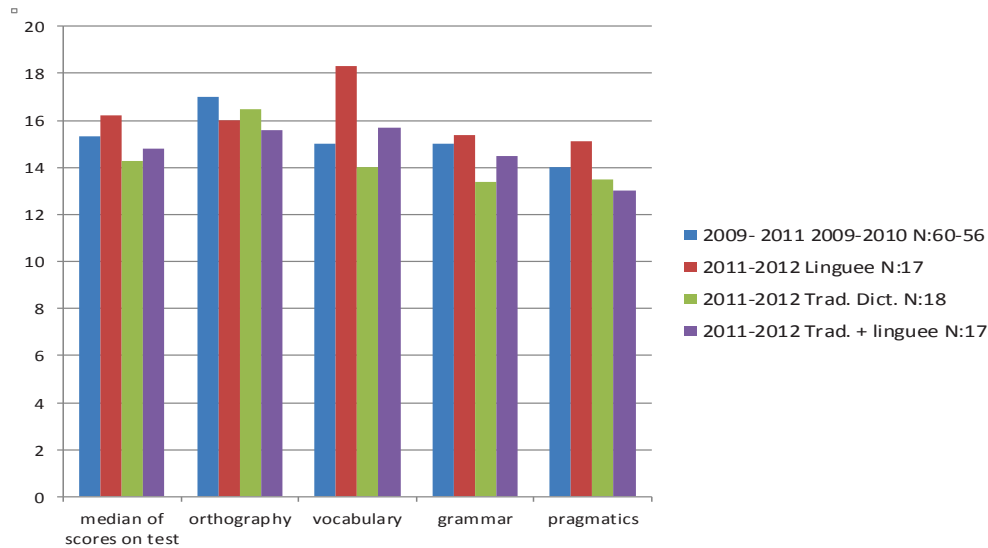


Figure 2. Results of the quantitative analysis.

Evidence for these results can be found in the qualitative analysis (Figure 3): CG2 is remarkably less satisfied with the time available for the test (more than 3 points lower, on a range of 5), although they show the highest satisfaction scores on the quality of the available tools (almost 1 point higher than EG). This seems to indicate that the combination of the available tools is higher, but too time-consuming.

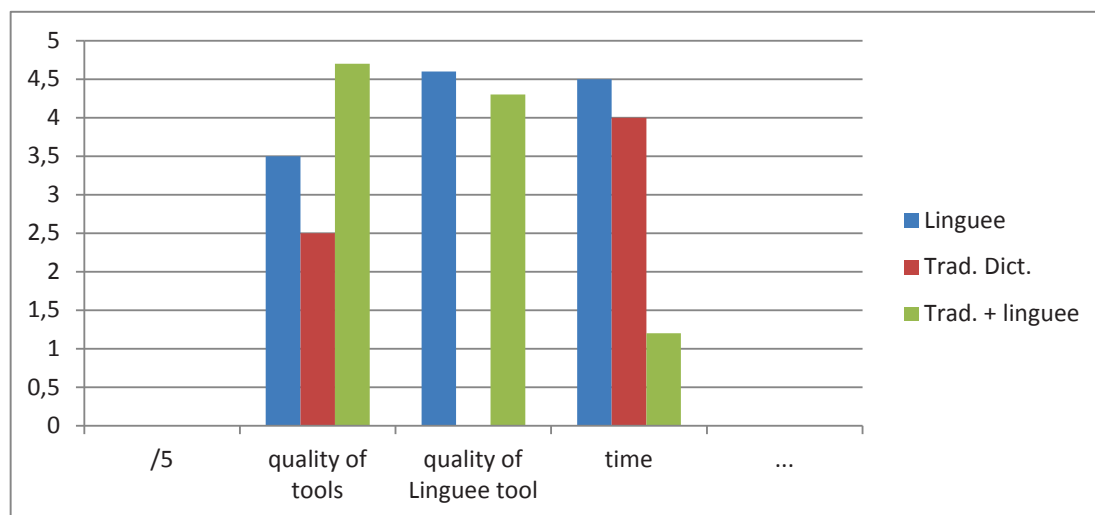


Figure 3. Results of the qualitative analysis.

4. Further pedagogical and technological interventions: further results?

As a result, it seems that a further integration of tools (such as spelling checkers, traditional dictionaries, corpora, data driven lexico-grammatical tools *in context*...) would offer a solution for CG2, as it would increase the usability of several tools which are very valuable *an sich*. An example of such an integrated tool, but for other languages than Spanish, is the French-English-Dutch *Interactive Language Toolbox* (<https://ilt.kuleuven.be/inlato/>). Nevertheless, an ongoing study (see Rymenams et alii, s.d.), comparing French (pre)tests, immediate posttests and delayed posttests of 107 KU Leuven users with a control group of 105 more KU Leuven students (both groups from Economical Sciences), resulting in 754 texts with an average of 319,96 words, concludes that the tool has above all a positive impact on the formal quality of the texts as well as the learning of French in the short run, but that for longer-term learning this effect is less clear and seems to be completely absent for the weaker students. However, these students were not given any instructions on the use of the tool.

This brings us back to the beginning of our article, where we insisted on the importance of “triggers and rewards”, facilitated by the teacher who fulfills the role of advisor or facilitator, as well as on the importance of “hands-on experiences” (demonstrations in the classroom, feedback, use of portfolios...). In other words, we propose a double intervention: a technological one, integrating several kinds of tools into one “toolbox”, and a pedagogical one, integrating the use of such a toolbox into the classroom.

References

- Bank, Christina (2010). Die Usability von Online-Wörterbüchern und elektronischen Sprachportalen. MA dissertation. University of Hildesheim.
- Bowker, Lynne & Jennifer Pearson. (2002). Working with Specialized Language. A Practical Guide to Using Corpora. London & New York: Routledge.
- Buyse, Kris. (2008). Analysis of the course materials of Economic Spanish of four institutions and of the interviews with their teachers. Leuven: unpublished raw data.
- Buyse, Kris. (2011). Effective writing tasks and feedback for the internet generation. *Language Learning in Higher Education*, 1 (2), 377-398.
- Buyse, Kris, Nicole Delbecque & Dirk Speelman. (2009). Trampas y pistas para la expresión escrita de los neerlandófonos. Averbode - Madrid: Averbode / Edelsa.
- Cervero, María Jesús & Francisca Pichardo Castro. (2000). Aprender y enseñar vocabulario. Madrid: Edelsa.
- Chambers, Angela (2007). Popularising corpus consultation by language learners and teachers. In E. Hidalgo Tenorio, L. Quereda Rodríguez-Navarro & Juan Santana (Eds.), *Corpora in the Foreign Language Classroom: Selected Papers from the Sixth International Conference on Teaching and Language Corpora (TaLC 6)*. University of Granada, Spain, 4-7 July, 2004.

- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Gijsselaers, Wim. (2007). *Talking about My Generation*. Keynote at EARLI 07: European Practice-based and Practitioner Research conference on Learning and Instruction for the new generation. Maastricht, 14-16 November.
- Gouws, Rufus. (2011). Learning, Unlearning and Innovation in the Planning of Electronic Dictionaries. In P.A. Fuertes-Olivera & H. Bergenholtz (Eds). *e-Lexicography. The internet, digital initiatives and lexicography* (pp. 17–29). London, New York: Continuum.
- Heid, Ulrich. (2011). Electronic Dictionaries as Tools: Toward an Assessment of Usability. In P.A. Fuertes-Olivera & H. Bergenholtz (Eds). *e-Lexicography. The internet, digital initiatives and lexicography* (pp. 287–304). London, New York: Continuum.
- Nielsen, Sandro & Lise Mourier. (2007). Design of a function-based internet accounting dictionary. In H. Gottlieb & J. E. Mogensen (Eds.), *Dictionary Visions, Research and Practice. Selected Papers from the 12th International Symposium on Lexicography* (pp. 119–139). Copenhagen 2004. Amsterdam/Philadelphia: John Benjamins.
- Nomdedeu Rull Serrano, Antonio (2009). *Diccionarios en Internet para el aula de ELE*. *RedELE: Revista Electrónica de Didáctica ELE*, 15, 2009.
- Pastor, Verónica & Alcina, Amparo. (2010). Search Techniques in Electronic Dictionaries: A Classification for Translators. *International Journal of Lexicography* 23.3, 307-354.
- Rymenams, Sarah, et alii (s.d.). *SOS français: Conception et évaluation d'un didacticiel d'aide à la rédaction interactif*. Leuven : s.e.
- Schmitt, Norbert. (2000). *Vocabulary in Language Teaching*. Cambridge: CUP.